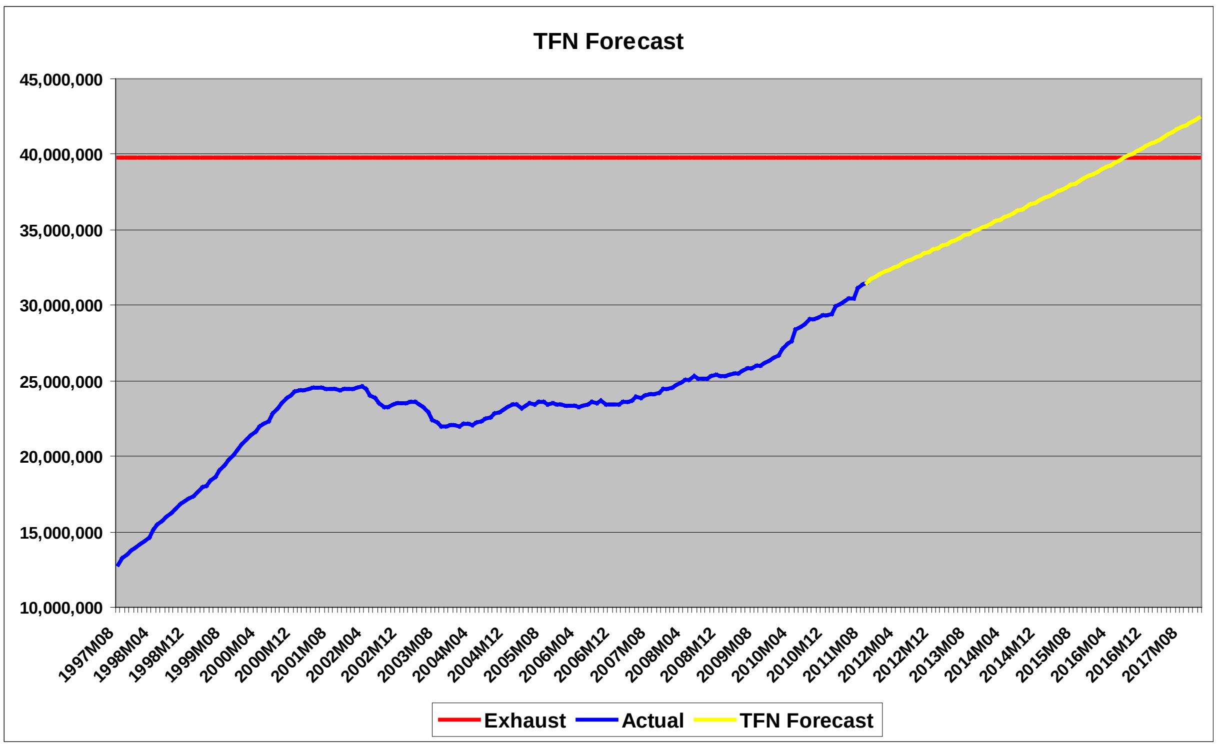
OffHook Consulting Inc.

Forecasting Utilization of Toll-Free Numbers in North America

Prepared for SMS/800, Inc.
November 2011

Executive Summary

In the first quarter of 2008, the prior method of assessing the exhaust date of toll-free number (TFN) capacity was abandoned in favor of using more formal statistical time series forecasting techniques.[1] Most recently, data beginning in August 1997 and running through the end of September 2011 was examined. OffHook considered a variety of candidate statistical models using an approach comparable to that employed in previous studies, the most recent in March 2010.[2] Relative to other candidates, the model selected had the best statistical characteristics, as well as a slightly earlier predicted exhaust date and somewhat higher assessment of risk of exhaust during preceding periods. Thus, the choice of this model is consistent with conservative, but statistically supported, decision making with respect to planning for the next code opening. The selected model's TFN forecasts (in yellow) and the TFN exhaust value (in red) are shown below.



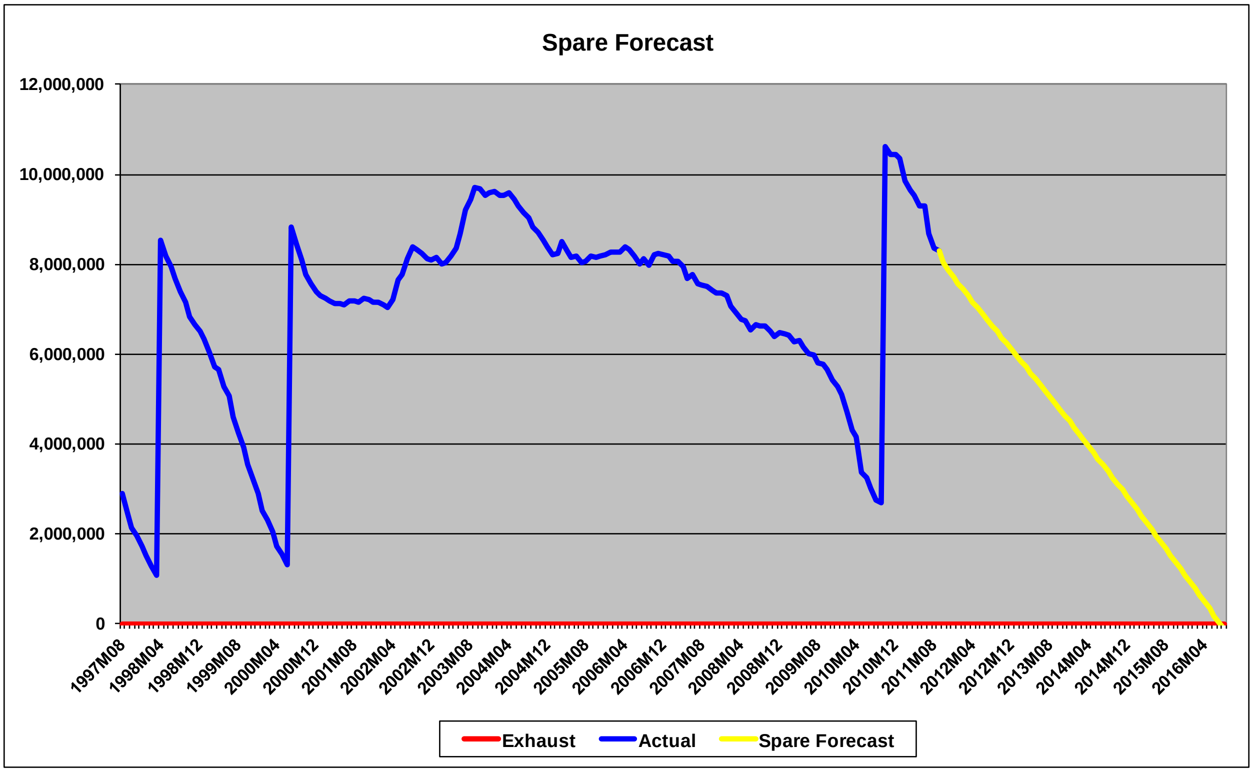
At the end of September 2011, 79% of the TFN pool was in use. While the forecast indicates exhaust of the currently available TFN pool in August 2016, the following table shows the forecasted dates for reaching 85%, 90% and 95% of the current TFN pool.

[1] The old method, used for more than a decade, employed a six-week rolling average of weekly number growth (or decline) and was highly volatile.
[2] Although TFN forecasts are planned every six months, SMS/800, Inc. decided to postpone the forecasts scheduled for 3Q2010 and 1Q2011 due to the 855 code opening in October 2010.

% of TFN Pool	TFNs	Date Reached
85%	33,821,578	Feb 2013
90%	35,811,083	April 2014
95%	37,800,587	June 2015

It is worth noting that the 855 code was opened in October 2010 with 91.5% of the TFN pool in use, or approximately 2.68 million TFNs available in the spare pool. In the year since the code opening, 2.36 million more TFNs are now in use, consuming nearly the entire spare pool at the time 855 was opened.

This information can be viewed in a different way; by showing a forecast of the stock of *spare* TFNs. As of the end of September 2011, there were approximately 8.31 million spare numbers available from the total pool of 39,790,092. Note the forecasted stock of spare TFNs drops below 2 million in June 2015.



The business decision facing SMS/800, Inc. is when to start the planning and development effort required to open a new toll-free Numbering Plan Area (NPA). Although forecasting an exhaust date is helpful in that process, it does not adequately convey the degree of risk and uncertainty (whether large or small and inherent in any forecast) as to when exhaust may occur. Thus, OffHook's forecast methodology also assesses the risk of exhaust at particular dates in the near future.

It is our understanding that the SMS/800 industry requires 18-24 months to prepare and open a new toll-free NPA, e.g., 844. The forecasting model indicates that the risk of exhausting the current stock of TFNs:

- Is less than 1% through November 2012
- Is less than 10% through August 2013
- Is 20% in April 2014, and
- Is less than 30% through December 2014

The risk is significant enough over the next two years (at 13.6% by the end of November 2013), that SMS/800 should continue monitoring the TFN stock vis-à-vis the forecasted TFN demand.

The statistical techniques used to develop these projections rely on the historical data and thus cannot predict turning points or dramatic changes in growth that are not implicit in the historical data. Thus, the forecasting exercise that OffHook has performed simply describes the most likely date that toll-free number capacity will be reached *if the data patterns of August 1997 through September 2011 should continue into the future*.[3]

OffHook believes that statistical analyses are the most appropriate methods for projecting the exhaust of the TFN pool and assessing the risk surrounding the business decisions. However, the industry has historically relied on a method published by ATIS in the "Toll Free Resource Exhaust Relief Planning Guidelines" (ATIS-0300057, July 1998) to determine when the TFN pool would exhaust and therefore this approach was also examined.

Using monthly and quarterly data, and assessing the results for multiple sample sizes, the predicted exhaust dates vary from the second quarter of 2014 to second quarter 2016. The approach, analysis and results are examined in more detail in Appendix A.

[3] Time series techniques model only patterns of motion in the data and cannot account for changes in underlying external forces. If forces underlying the motion of the series change, then time series techniques must change models to adapt to new patterns in the data.

Introduction

SMS/800 has a limited supply of toll-free numbers (TFNs). There are currently five open Numbering Plan Areas (NPAs): 800, 866, 877, 888, and 855. At the end of September 2011 31,475,626 (79%) of the available 39,790,092 TFNs were in use, leaving approximately 8,314,000 spare numbers. Based on our understanding, the toll-free industry requires approximately 18-24 months of lead time to complete the appropriate system changes. Accordingly, SMS/800 has an ongoing forecast effort to assess when, or even if, the available numbers will exhaust. The objectives of the forecast of Toll-Free Numbers in Use (TFNs) are to: a) estimate a timeframe that the currently open TFN number pool will exhaust; and, b) assess the risk that the TFN number pool will exhaust over the next one to two years.

Prior to the first quarter of 2008, the method used to assess when toll-free number capacity would exhaust was to divide the quantity of spare numbers by the most recent six-week rolling average of weekly number growth. This approach produced weekly results that were highly volatile. The objective of the updated forecast method is to produce a forecast that: (a) does not fluctuate substantially with relatively minor changes in recent values of toll-free number usage; (b) reflects the uncertainty inherent to any forecast by including a range of results; (c) is based on long-term historical patterns in toll-free number usage rather than short-term perturbations; and (d) is generated by a defensible empirical methodology.

After examining alternative forecasting methodologies, including econometric models, SMS/800, in April 2008, adopted a purely statistical time series forecasting approach. This approach relies on past values of toll-free number use and considers possible changes in the rate of increase over time; it generates toll-free number forecasts as well as measures of the uncertainty surrounding these forecasts at particular decision dates in the future. Rather than relying on a single statistical model, this approach considers multiple statistical models from which the best models are selected.

The TFN forecast here is consistent with, and builds upon, the work performed in previous TFN forecasts, as well as what was learned in the 2009 CRA forecasting methods analysis.[4]

Forecasting Approach and Basic Methodology

OffHook utilized TFN data available for the period 8/23/1997 through 9/24/2011 on a weekly-ending basis (737 weeks). The weekly data was converted to monthly data and quarterly data and also translated (in some instances) into first differences and natural logs.[5]

[4] In 2009 OffHook performed research related to forecasting methods for the CRA (as well as TFN) and produced the report “CRA Forecasting Methods Analysis”.
[5] The data is described in more detail in a separate section.

As with recent (over the last two-plus years) work in both CRA modeling and TFN forecasting, our success in using ARIMA models caused us to consider these models again for forecasting TFNs.[6] In addition, our work in March 2010, demonstrated that Autoregressive Conditional Heteroskedasticity (ARCH/GARCH) modeling techniques could be useful to address fluctuations in the variance in the error terms that tend to arise with the use of longer sample periods. OffHook again considered ARCH/GARCH techniques and full sample periods.[7]

OffHook's efforts did not specifically attempt to employ economic data or econometric approaches to forecasting TFNs. Such approaches likely would require significantly more effort, and may well not produce results superior to purely statistical time series approaches. While such approaches may be worth pursuing in the future, OffHook does not believe that they are currently warranted.

It should be noted that statistical time-series methods, while generally sound and defensible, essentially assume that past behavior is a window into the future. They implicitly assume that TFN usage will exhibit essentially the same behavior in the future as it has over the historical period used to estimate the model. If historical TFN usage were increasing rapidly, then the forecast should show continued rapid growth. Alternatively, if historical TFN were growing gradually, then forecasted TFN should generally exhibit the same behavior.

Both the monthly and quarterly data sets were bifurcated into an estimation period and a hold-out period. The latter is used for model selection to test the forecast accuracy of models built using the estimation period. Use of a two-year hold-out period is preferable since this corresponds to the outer range of lead time to implement a new toll free code. However, starting in late 2009, TFN growth accelerated substantially. In order for the estimation sample to include some of this period (and for the final forecast model to be trained on this period of recent accelerated growth), the current analysis focused on hold-out periods of 18 and 12 months. Future updates to this analysis in 2012, for example, may permit a return to the use of the preferred two-year hold-out period for model selection.

---

[6] An ARIMA model, summarized by ARIMA (p,d,q), can be characterized by three categories of parameters: **p,** the longest number of months by which past data directly influence current data, also referred to as the autoregressive (AR) term; **d,** the number of times the series (i.e. TFN) is differenced to recognize the degree of increase or decrease over time; and **q,** the longest number of months by which lagged forecast errors improve the prediction of current data. The lagged forecast error **q** term is also referred to as a "moving average" (MA) term; this term is akin to creating an exponentially weighted average of past data (of TFN or its degree of increase or decrease in this instance), with the most recent data given the highest weight and the weights assigned to older data exhibiting exponential decay.

[7] GARCH and ARCH terms, described later.

Various statistical model specifications were estimated with data from the estimation periods.[8] The forecasts from these models were then compared to the actual TFN values for the corresponding hold-out period. Better performing models were re-estimated with the full sample of data (through September 2011 for the monthly models and through 2011Q3 for quarterly models).

The candidate models were assessed on the basis of the following characteristics:

1) statistical significance of the modeled terms;
2) whether the resulting in-sample residuals appeared to be random (i.e., exhibited white noise);
3) the principle of parsimony (the fewest terms employed to still fit the data and produce white noise);
4) errors in the hold-out period (Mean Percent Error, MPE, and Mean Absolute Percent Error, MAPE, and Theil measure of forecast bias);
5) the size of the confidence interval around the estimate;
6) robustness – whether the terms changed or their coefficients changed significantly when the models were re-estimated with the full data set; and,
7) whether the forecasts made sense, based upon our knowledge of TFNs.

As noted above, the business decision facing the SMS/800 is when to start the planning and development effort required to open a new toll-free Numbering Plan Area (NPA). As an extension of the original base forecasting efforts, OffHook has expanded the focus to assess, more explicitly, the risk of exhaust over relatively near-term (e.g., 2 year) periods. These values are more germane to the business decision of starting the process to open a new toll-free NPA. This risk assessment utilizes information about the variance underlying the model forecasts (similar to the information used in creating a confidence interval about an estimate). This is described in more detail in a later section of the report.

The Data Set and Sample Period

Weekly historical data for TFNs in use are available starting in August 1997 through September 2011 (737 weeks in total).[9] A monthly TFN series (168 months) was created by taking the last weekly value within the month. A quarterly TFN series (56 quarters) was created by taking the last weekly value within the quarter.

---

[8] The specific data for the estimation period is determined by the length of the hold-out period. For example, with an 18 month hold-out period and monthly data, the hold out period was April 2010 to September 2011, while the estimation period was August 1997 to March 2010.

[9] Several "gaps" in the weekly series prior to 2002 were filled in using a simple interpolation method. All data provided were assumed to be accurate. No cleansing of raw data was done to correct potential typos or other errors.

Statistical examination of the data series reveals a notable characteristic of the TFN series - the variation or fluctuation in TFN over time.[10] Specifically, the pre-2001 period is characterized by a much larger degree of variation in TFN than the post-2001 period. Moreover, as seen in the table below, the variation across data frequencies (i.e., weekly, monthly, quarterly) is essentially the same for each period examined.

In contrast, when the three data series are transformed (from levels) into first differences to reflect TFN's rate of increase or decrease from one period to the next, these conclusions change. As shown in the figure below, the degree of variation across the series differs by the periodicity of the data with the longer periodicity exhibiting lower variation (i.e., first-differenced quarterly series exhibits the least variation across the full sample, and the weekly first-difference series exhibits the most variation across the full sample). Thus, models estimating the rate of change in TFN are likely to perform better using monthly or quarterly data rather than weekly data.

		Coefficient of Variation				
		1997Q4 - 2000Q4	2001Q1 - 2003Q4	2004Q1 - 2009Q3	2009Q4 - 2011Q3	1997Q4 - 2011Q3
Weekly	**Level**	17.43%	3.51%	4.47%	5.54%	15.51%
Monthly	**Level**	17.58%	3.54%	4.49%	5.58%	15.53%
Quarterly	**Level**	18.01%	3.60%	4.55%	5.78%	15.61%
Weekly	**Difference**	62.45%	-344.56%	407.50%	124.71%	238.17%
Monthly	**Difference**	29.66%	-263.92%	169.23%	59.79%	155.84%
Quarterly	**Difference**	17.72%	-236.13%	102.41%	40.17%	136.54%

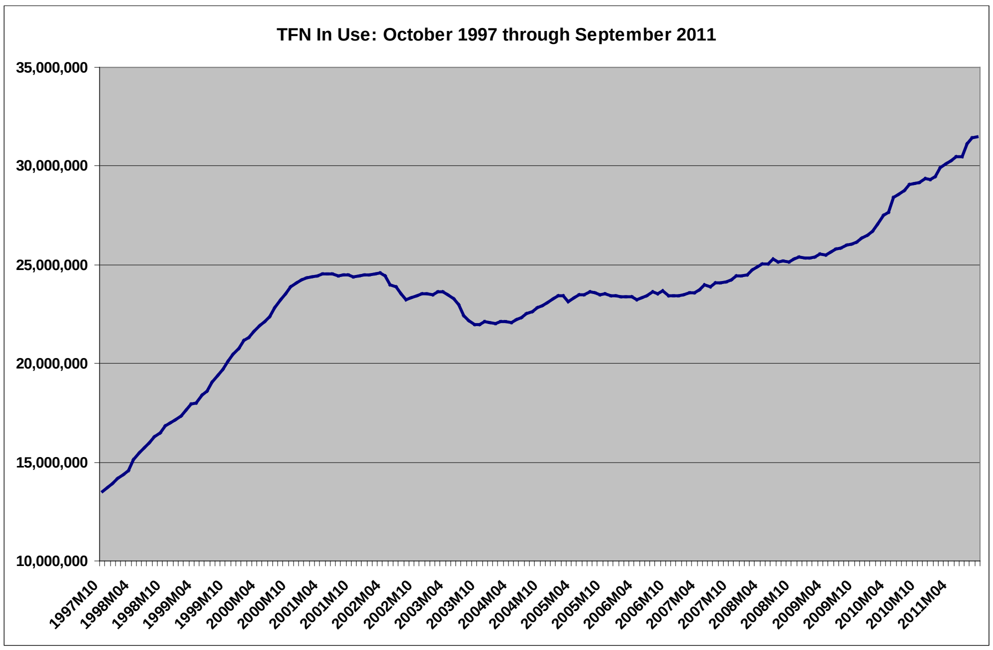
Also note that, regardless of the periodicity of the data, the post-2001 period exhibits a much higher degree of variation (in first differences) than the pre-2001 period (opposite of the pattern in the un-differenced, or levels, series). Differenced TFN shows greater variation after 2001, especially 2001-2003, a period characterized by a general decline in the level of TFNs and several turning points. Thus, models estimating the rate of change in TFN may perform better if the model accounts for changes in variance over time. ARCH/GARCH modeling is one approach to incorporating changes in variance.

In the past, there had been some concern about sufficiency of sample size when using quarterly data. A statistical rule of thumb is to employ samples with 40 or more observations for ARIMA modeling. With an 18-month hold out sample, the estimation sample (when using quarterly data) is 50 quarters, and 56 quarters for the re-estimation process; these samples are sufficiently large for the application of ARIMA analysis. The table below identifies the quantity of observations for each period.

[10] In statistical terms, TFN data are non-stationary, which suggests that it may be more appropriate to model rates of change (i.e., first differences) rather than the level of TFN.

		Quantity of Observations				
		1997Q4 - 2000Q4	2001Q1 - 2003Q4	2004Q1 - 2009Q3	2009Q4 - 2011Q3	1997Q4 - 2011Q3
Weekly	**Level**	170	156	300	105	731
Monthly	**Level**	39	36	69	24	168
Quarterly	**Level**	13	12	23	8	56

One critical question is whether the entire period for which historical data exists is germane to the current forecast for TFN usage. As can be seen from the graph below, there appear to be significant variations in the growth patterns of TFNs over the past fourteen years.



The use of the full sample period, while providing more data, means that first difference models are estimated across a data set in which the coefficient of variation in the data changes significantly across time periods; this has the potential to produce heteroskedasticity in the estimated residuals of the models. To address this issue, (i.e. to improve the statistical efficiency of the long-run forecasts) ARCH/GARCH terms[11] are often employed in the modeling process.[12] From work performed in March 2010,

[11] An autoregressive conditional heteroskedasticity (ARCH) modeling term considers the variance of the current error term to be a function of the variances of the previous time periods' error terms. It is employed commonly in modeling financial time series that exhibit time-varying volatility clustering. If an autoregressive moving average model is assumed for the error variance, the model is a generalized autoregressive conditional heteroskedasticity (GARCH). See, e.g., http://en.wikipedia.org/wiki/Autoregressive_conditional_heteroskedasticity.
[12] GARCH addresses the residual variance in TFN still unexplained after the ARIMA modeling. For example, the weighted lagged squared residual represents "news" about the evolving variance.

OffHook found that the introduction of ARCH/GARCH terms may allow the entire sample (back to August 1997) to be used and better model the time series. In the analysis, OffHook conducted hold-out tests of forecasts from full-sample models versus models estimated from 2004 to present. The analysis found that forecasts from models estimated over the full sample period, with the possible addition of GARCH terms, yielded better statistical results. Moreover, use of the full sample insures sufficient sample size for quarterly models to be considered.

Model Estimation and Selection

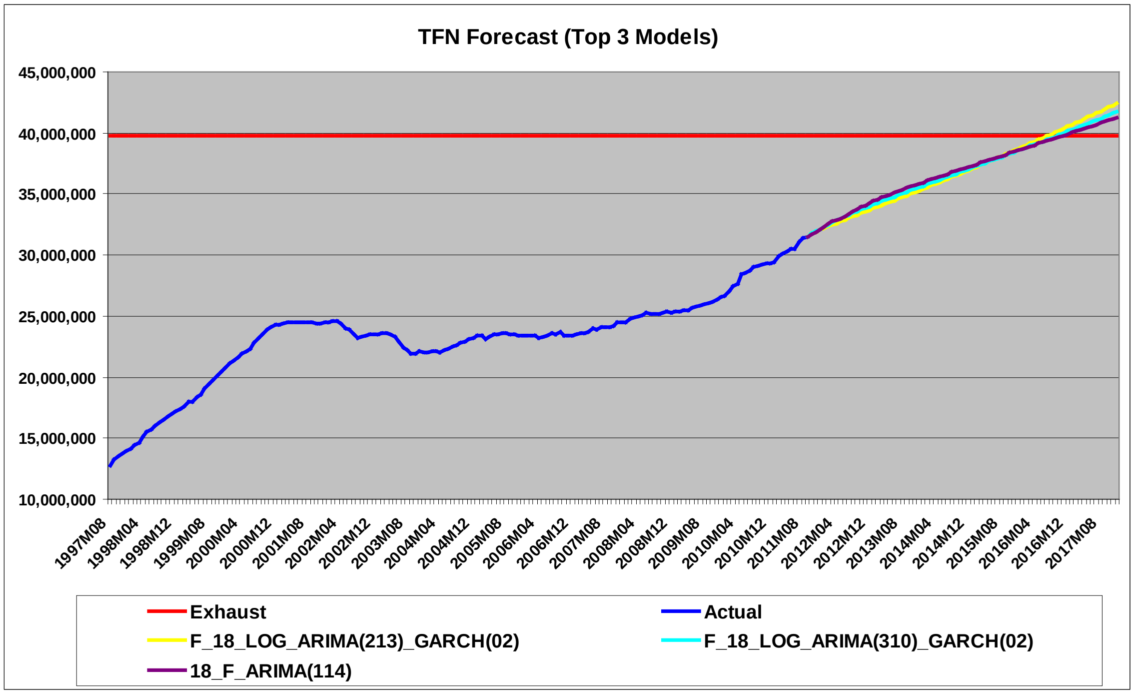
For each sample period (and corresponding hold-out period) OffHook estimated several specific types of models. Several models had relatively small errors (less than 1% MAPE) for the 12 month and/or 18 month hold-out period. Because several of the models produced small absolute errors, other criteria guided the model selection process (as listed in the approach and methodology section above).

Once the top candidate models had been selected, the models were re-estimated, using the full sample (with data through 2011M09 for monthly models and 2011Q3 for quarterly models). The re-estimated models were evaluated again, adding consideration of robustness of the model terms and their estimated coefficients and the reasonableness of the forecast results.

While several models had acceptable statistical characteristics, the list of candidate models was ultimately narrowed to three “finalists”; one estimated with quarterly data and two with monthly data. Each of the three models utilizes first difference data, which provided superior statistical performance to the use of the stock TFNs. This aspect of the statistical analysis points out the importance of focusing on the rate of change of the series over time when generating a long-term forecast.

	F_18_LOG_ARIMA(213) GARCH(02)	**F_18_LOG_ARIMA(310) GARCH(02)**	**18_F_ARIMA(114)**
Frequency	Month	Month	Quarter
Dep Var (TFN)	Log	Log	Non-Log
Differencing	1x	1x	1x
Constant	Yes	Yes	Yes
AR Terms	1,2	1,2,3	1
MA Terms	3	0	4
ARCH Terms	0	0	0
GARCH Terms	1,2	1,2	0
18-Month MAPE	0.70%	0.77%	0.70%
18-Month MPE	0.19%	0.30%	-0.19%
18-Month Theil Bias	0.056	0.116	0.06
Exhaust Date	August 2016	October 2016	4Q 2016
Date of 20% Risk	March 2014	February 2014	4Q 2014

The two monthly models employ a natural log specification and ARCH/GARCH terms, but the quarterly model does not. Use of a natural log as found in the monthly forecasts implies that TFN growth compounds over time in percentage terms. The quarterly forecast embodies no such compounding.[13] The table above provides a description of the three finalist models. Note that each model has very low mean absolute percentage error (MAPE) for the 18-month holdout period (less than 1%), and low bias. A graph of the forecasts produced by each of the three finalist models, below, shows that the three models produce similar forecasts.



At this point, there is little to distinguish among the three models; each has low errors and bias, and all produce similar forecasts. Ultimately the log monthly model ARIMA (2,1,3) GARCH (0,2) was selected since it produces the earliest exhaust date (and hence is the most "conservative") and because technically it has marginally smaller bias and error.

Risk Assessment

As noted above, OffHook augments the point forecasts to assess the risk of exhausting the TFN number pool sooner than the point forecast. These risk values are germane to the business decision of when to start the process to open a new toll-free code.

[13] One may then ask: why are the forecasts so strikingly similar? The rate of growth of TFNs is sufficiently small over the forecast period that compounding is not substantial, so the monthly logarithmic forecasts and quarterly forecasts yield almost identical exhaust dates.

Specifically risk is defined here as the probability that TFN demand exceeds the stock of TFNs available under existing codes. For example, risk of 25% on a specific date suggests that there is a one in four chance that TFNs will exhaust by that date or 3:1 odds against exhaust. Risk of 50% on a specific date suggests there are even odds that any spare TFNs remain available. This is the most probable date of exhaust. After that date, the odds are higher than 1:1 that TFN demand will exceed the stock of available numbers.

The measure of risk relies upon the measure of the standard error of the forecasts and is conceptually similar to a confidence interval.[14] The table below includes the risk at one year intervals (from the release of this report),

November of (Year)	Exhaust Risk
2011	0.0%
2012	1.3%
2013	14.9%
2014	29.6%
2015	42.4%
2016	53.1%

Although the model indicates the most likely date of exhaust is nearly 5 years away, the risk of exhausting earlier exceeds 25% within 3 years. Given that the industry requires at least two years to prepare for the opening of a new code, these risk levels suggest that careful monitoring of TFN demand may be warranted.

Some Cautions and Recommendations

The statistical techniques used to develop these projections rely on the historical data and thus cannot predict turning points or dramatic changes in growth that are not implicit in the historical data. Thus, the forecasting exercise that OffHook has performed simply describes the most likely date that toll-free number capacity will be reached *if the data patterns of August 1997 through September 2011 were to continue into the future*.

[14] Measured as, Risk = 1 – *N((*Exhaust TFN – Forecast TFN)/Forecast Standard Error) where *N* is the cumulative standard normal distribution.

Appendix A – ATIS Approach to Determining an Exhaust Date for the TFN Pool

OffHook believes that statistical analyses such as those described within the body of this report, are the most appropriate methods for projecting the exhaust of the TFN pool. Further, OffHook believes that statistical methods have the ability to assess risk surrounding the business decisions related to the exhaustion of the TFN pool, e.g., the risk that the TFN pool will exhaust sooner than the point estimate. However, the industry has historically relied on a method developed by ATIS to identify the date at which the industry should be notified of an impending code opening. This approach is described in the "Toll Free Resource Exhaust Relief Planning Guidelines" (ATIS-0300057), published in July 1998.

Exhaust of the TFN pool is determined by inputs for average demand and accelerated demand to be provided by SMS/800. The prescribed steps to determine the predicted number of months remaining until exhaust are represented by the following formula:

(Spares-(accelerated demand * # months of accelerated demand))/(average demand) + # of months of accelerated demand

The approach is straightforward and simple to implement making the approach appealing. However, as these inputs are not well defined, the approach is prone to be arbitrary and can lead to highly volatile results based on different input choices. Further, there may not be an obvious rationale by which to choose between input values. Still, it is OffHook's understanding that this remains the official method for determining when the toll-free industry should be notified of a code opening, and thus a version of the method was implemented.

OffHook's variation on this approach relies upon the same formula, but includes more narrowly defined inputs, as follows:

- The average demand is defined as the average of all data points (within the sample size selected) within a range of the overall average plus and minus the standard deviation;
- Accelerated demand is defined as the average of all data points (within the sample size selected) greater than the overall average plus the standard deviation; and,
- The quantity of periods for accelerated demand is the quantity of data points in the range greater than the overall average plus the standard deviation.

An attempt to examine many alternatives using the ATIS approach led OffHook to prepare this analysis using monthly and quarterly data, and to assess the results for many sample sizes. The following tables for monthly and quarterly samples are demonstrative of the results.

As with many analyses of data sets, the period over which the data is considered has a significant impact on the results. For the purposes of this exercise, OffHook examined the most recent four years of toll free number data, and based the ATIS approach on this data set. The following two tables represent the results based on the analyses performed on months and quarters.

Figure A: ATIS Exhaust Approach - Months

# of Months in Sample	6	12	18	24	36	48
Spares	8,314,466	8,314,466	8,314,466	8,314,466	8,314,466	8,314,466
Average Demand	182,821	156,622	160,873	186,165	132,240	118,580
Acc Demand	622,446	558,795	635,663	635,663	538,361	538,361
Months of Acc Demand	1	2	3	3	5	5
Months to Exhaust	43.00	48.00	43.00	37.00	48.00	52.00
Exhaust Date	Mar-15	Aug-15	Mar-15	Sep-14	Aug-15	Dec-15

Figure B: ATIS Exhaust Approach - Quarters

# of Quarters in Sample	2	4	8	12	16	20
Spares	8,314,466	8,314,466	8,314,466	8,314,466	8,314,466	8,314,466
Average Demand	675,570	571,597	758,720	391,473	370,296	310,338
Acc Demand	NA	994,140	1,060,295	1,004,813	1,004,813	950,158
Quarters of Acc Demand	NA	1	1	3	3	4
Quarters to Exhaust	12	14	11	17	17	19
Exhaust Date	2014Q3	2015Q1	2014Q2	2015Q4	2015Q4	2016Q2

As seen in the figures above, the ATIS approach indicates exhaust between the second quarter in 2014 and second quarter 2016.